Sparsity-based Autofocus for Under-sampled Synthetic Aperture Radar

Sparsity-based Autofocus for Under-sampled Synthetic Aperture Radar

Shaun I. Kelly, *Student Member, IEEE,* Mehrdad Yaghoobi, *Member, IEEE,* and Mike E. Davies, *Senior Member, IEEE,*

***Abstract*—Motivated by the field of compressed sensing and sparse recovery, nonlinear algorithms have been proposed for the reconstruction of synthetic aperture radar images when the phase history is under-sampled. These algorithms assume exact knowledge of the system acquisition model. In this paper we investigate the effects of acquisition model phase errors when the phase history is under-sampled. We show that the standard methods of autofocus, which are used as a post-processing step on the reconstructed image, are typically not suitable. Instead of applying autofocus as a post-processor, we propose an algorithm that corrects phase errors during the image reconstruction. The performance of the algorithm is investigated quantitatively and qualitatively through numerical simulations on two practical scenarios where the phase histories contains phase errors and are under-sampled.**

***Index Terms*—Synthetic Aperture Radar, Autofocus, Compressed Sensing, Sparse Recovery, Blind Calibration, Block Relaxation Methods, Phase Retrieval**

I. Introduction

SYNTHETIC aperture radar (SAR) is an active ground imaging system which is based on the coherent processing of multiple radar echoes. Typically, the reconstructed image is formed from the stored echos (phase history) using a linear approximation of the pseudo-inverse, e.g. polar format algorithm (PFA), range migration algorithm (RMA) or filtered back-projection. The approximate pseudo-inverse is an inverse which is defined on a finite region of the spatial frequency support of the reconstructed complex image. The size of this support is defined by the transmitted RF signal bandwidth and the size of the synthetic aperture. Ideally the reconstructed image would have a rectangular support in the spatial frequency plane so that the point spread function (PSF) would be a two-dimensional sinc function. This is approximately the case in systems where the synthetic aperture is uniformly sampled and the transmitted RF signal has a contiguous bandwidth. However, in a number of interesting non-standard SAR scenarios, this will not be true.

Two such systems that we will consider in this paper are multifunction and ultra wide band (UWB) SAR. In a multifunction SAR system, the radar antenna is used for multiple tasks which causes interruptions in the uniform acquisition of SAR data along the synthetic aperture [1], [2]. In the case of UWB SAR, the transmitted signal spectrum is broad and may contain frequency sub-bands that are in use by other communication systems or where transmission is not allowed. To avoid interference, notch filters are commonly used in the transmitter and/or the receiver to avoid using these sub-bands [3], [4].

In both of these scenarios defining an inverse on an approximately rectangular spatial frequency support is ill-posed. Fig. 1 demonstrates why a rectangular support is sought by comparing the PSF of a rectangular spatial Fourier support and a randomly under-sampled aperture. In the PSF of the under-sampled aperture, unlike the ideal PSF, a significant amount of the target energy is contained in the side-lobes. Clearly this is undesirable. In order to make this problem well-posed, an appealing idea is to apply the tools and theory of compressed sensing (CS) and sparse recovery, for example [5], [6], [7], [8].

The theoretical results of CS are based on exact knowledge of the linear acquisition system, however, in practical situations, such a system cannot be perfectly known. This is the case in SAR where the received phase history may contain significant phase errors due to imperfect system modelling. Methods for correcting these errors in fully-sampled systems are known as autofocus algorithms and are most commonly used as a post-processing method on the reconstructed image.

All autofocus algorithms require a signal model for either the phase errors and the image or both. Additionally, many algorithms make a far-field and small aperture angle approximation so that the phase errors are constant along the range axis of the reconstructed image. One of the earliest autofocus algorithms to be developed was the mapdrift (MD) algorithm [9]. MD estimates the phase errors based on a low-order polynomial model for the phase errors along the cross-range direction. Phase gradient autofocus (PGA), one of the most commonly used algorithms, requires the phase errors along the cross-range direction to vary smoothly and also requires the image to contain isolated point scatterers [10]. Recently another algorithm, multichannel autofocus (MCA), has been proposed which requires the focused image to contain a known region which is almost zero [11]. Although these post-processing autofocus methods have been very successful for correcting phase errors in fully-sampled scenarios, they may not be suitable for under-sampled SAR.

The algorithm proposed in this paper for image reconstruction and autofocus of a under-sampled phase history has similarities with the proposed method in [12]. Although the method proposed in [12] primarily concentrates on the fully-sampled scenario it does demonstrate that it is also applicable to the under-sampled scenario. Both methods involve approximately solving the same non-convex problem but our algorithm has

The authors are with the Institute for Digital Communications (IDCoM), University of Edinburgh, UK, EH9 3JL, e-mail: shaun.kelly@ed.ac.uk.

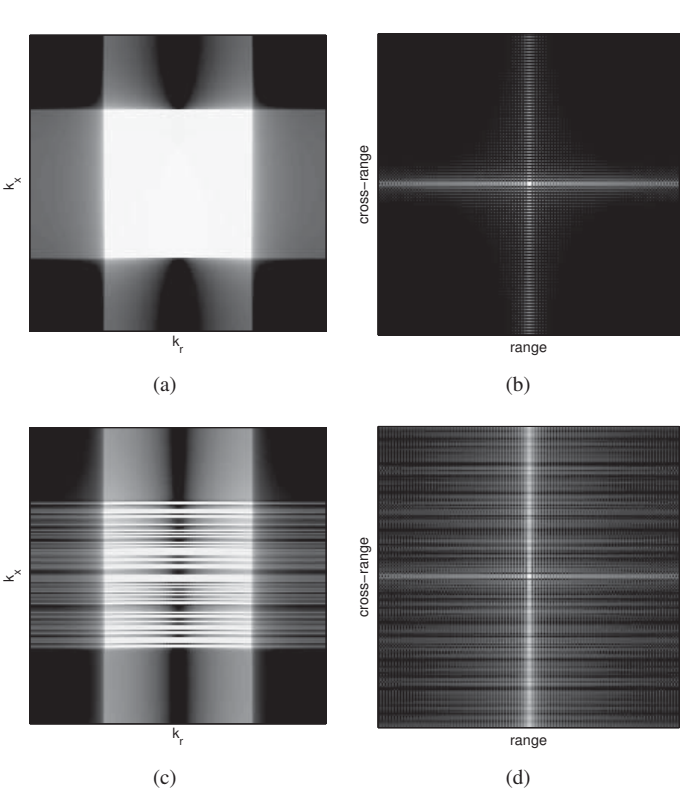


Fig. 1. PSF for full-sampled and under-sampled almost rectangular spatial Fourier supports: (a) is the fully-sampled support and its PSF is given in (b). (c) is the under-sampled support and its PSF is given in (d).

some additional practical benefits. Firstly, it can be shown to be stable and it produces a sequence that convergences to a connected set. Secondly, it empirically converges in a significantly smaller number of iterations.

A closely related problem which has been investigated in the signal processing literature in the last few years is the problem of phase retrieval, e.g. [13] and [14]. The goal of phase retrieval is to recover a complex signal from magnitude only measurements. The SAR imaging and autofocus problem is equivalent to the phase retrieval problem if we ignore all phase information due to a belief that it is corrupted. In these papers a technique known as "phase-lifting" is used to pose a convex problem which is solved to recover the signal. This technique involves "lifting" the signal so instead of recovering $\boldsymbol{x} \in \mathbb{C}^N$ the algorithm recovers $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^{\mathrm{H}} \in \mathbb{C}^{N \times N}$. This process is likely to be very costly computationally and will likely make these techniques infeasible for SAR systems.

Contributions of the paper

The main contributions of this paper are as follows. We show using CS theory and numerical simulations that standard post-processing autofocus methods are unsuitable for under-sampled SAR. We analyse under what conditions the image reconstruction and autofocus problem is well-posed. Also, we propose a new algorithm that correct phase errors within the image reconstruction algorithm. Empirically, we show that that this algorithm converges faster than existing methods and then, theoretically, we show that it is stable and convergent, which cannot be said of the existing algorithms. We also verify the performance of the algorithm using two practical scenarios.

Organisation of the paper

In Section II a brief background on relevant CS results is provided. Then in section III a SAR acquisition model is developed which includes phase errors. In Section IV the expected performance of existing post-processing auto-focus methods in a CS framework is investigated. Inherent ambiguities in the under-sampled phase error problem are analysed in Section V. A reconstruction algorithm for under-sampled SAR with phase errors is proposed in Section VI. Finally experimental simulations in Section VII are used to demonstrate the effectiveness of the proposed algorithm.

Notation

The following is a description of the notational conventions used within this paper. Matrices and vectors will be denoted by upper and lower case boldface symbols respectively (e.g. $\boldsymbol{X}$ and $\boldsymbol{x}$). Elements of matrices and vectors will be lower case lightface and will be indexed by subscripts, e.g. the element in the mth row and the nth column of a matrix $\boldsymbol{X}$ is denoted by x_{mn}.

The complex conjugate of a complex scalar x will be x^* and the complex conjugate transpose of a vector or a matrix will have a superscript H, e.g. $\boldsymbol{X}^{\mathrm{H}}$. $\boldsymbol{X}^{\dagger}$ is used to indicate the pseudo inverse of $\boldsymbol{X}$.

We define $\mathrm{diag}\{\boldsymbol{x}\}$ to be a square matrix with the elements of the vector $\boldsymbol{x}$ along its main diagonal. $\mathrm{Re}\{x\}$ will denote the real part of a complex scalar x.

The following notation is for matrix and vector norms. $\|.\|_0$ denotes the "counting norm" which is equal to the number of non-zero elements in a vector or a matrix. $\|.\|_{\mathrm{F}}$ and $\|.\|_1$ are element-wise two and one matrix norms respectively. Finally, $\|h\|_2 = \sup\{\|h(\boldsymbol{X})\|_{\mathrm{F}} : \|\boldsymbol{X}\|_{\mathrm{F}} = 1\}$ is the operator norm of a linear operator h.

II. Compressed Sensing: Background

CS theory provides a theoretical framework which can be used to analyse the reconstruction performance of an under-determined linear system, e.g.

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{n},$$

where, $\boldsymbol{y} \in \mathbb{C}^M$ are the measurements, $\boldsymbol{A} \in \mathbb{C}^{M \times N}$ is the system model, $\boldsymbol{x} \in \mathbb{C}^N$ is the original signal and $\boldsymbol{n} \in \mathbb{C}^M$ is a complex Gaussian noise for $M < N$. Without any further information, the best approximation of $\boldsymbol{x}$, in the MMSE sense, is given by the pseudo inverse $\boldsymbol{A}^{\dagger}\boldsymbol{y}$. However, using the tools of CS we may be able to produce a better estimate if $\boldsymbol{x}$ is sparse or well approximated by a sparse signal in an orthonormal basis, i.e.

$$\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{\alpha},$$

where, $\boldsymbol{\Psi} \in \mathbb{C}^{N \times N}$ is an orthogonal basis and $\boldsymbol{\alpha} \in \mathbb{C}^N$ is either a sparse vector, i.e. $\|\boldsymbol{\alpha}\|_0 \leq K$ for $K \ll N$, or is close

to its best K-term approximation $\boldsymbol{\alpha}_K$, i.e. $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}_K\|_2 \approx 0$. As well as the sparsity conditions on $\boldsymbol{x}$ we also require certain conditions on the matrix $\boldsymbol{A\Psi}$ to hold. A property that is commonly used to define sufficient conditions on $\boldsymbol{A\Psi}$ is the so-called restricted isometry property (RIP) [15]. A matrix $\boldsymbol{A}$ satisfies the (symmetric) RIP of order K if, for all vectors x with no more than K non-zero entries, there exists a (symmetric) RIP positive constant $\delta_K < 1$ which satisfies the following inequalities:

$$(1 - \delta_K) \|\boldsymbol{x}\|_2^2 \leq \left\| \hat{\boldsymbol{A}} \boldsymbol{x} \right\|_2^2 \leq (1 + \delta_K) \|\boldsymbol{x}\|_2^2 . \tag{1}$$

If $\boldsymbol{x}$ satisfies the sparsity condition and $\boldsymbol{A\Psi}$ satisfies a $2K$ order (symmetric) RIP with an RIP constant that satisfies the following inequality:

$$\delta_{2K} < \frac{3}{4 + \sqrt{6}} \approx 0.46515, \tag{2}$$

then $\boldsymbol{x}$ can be stably reconstructed from $\boldsymbol{y}$ using the following convex optimisation program [16]:

$$\begin{aligned} \underset{\boldsymbol{\alpha}}{\text{minimise}} \quad & \|\boldsymbol{\alpha}\|_1 \\ \text{subject to} \quad & \|\boldsymbol{y} - \boldsymbol{A\Psi\alpha}\|_2 \leq \sigma. \end{aligned} \tag{3}$$

The solution of Eq. (3), $\tilde{\boldsymbol{\alpha}}$, will be stable in the following sense

$$\|\tilde{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\|_2 \leq C_{1,K}\sigma + C_{2,K} \frac{\|\boldsymbol{\alpha} - \boldsymbol{\alpha}_K\|_1}{\sqrt{K}}, \tag{4}$$

where, $\sigma = \|\boldsymbol{n}\|_2$ and $C_{1,K}$ and $C_{2,K}$ are constants [17]. In words, our solution will be bounded by something that is proportional to the noise energy σ and the error associated with the best K-term approximation of $\boldsymbol{\alpha}$.

Although there is no computationally efficient way to check Eq. (2) for arbitrary matrices there are interesting asymptotic results for random matrices. One such result [18] is, if $\boldsymbol{A}$ is formed from $M < N$ random columns of a Fourier matrix and $\boldsymbol{\Psi}$ is an identity matrix then with overwhelming probability the matrix $\boldsymbol{A}$ satisfies Eq. (2) if M is of the order

$$M = \mathcal{O}(K \log^5(N)) \tag{5}$$

This result motives the use of CS theory for under-sampled SAR. If the under-sampled SAR observation matrix is similar to a randomly under-sampled Fourier matrix and our image contains only a small number of bright targets in clutter we may be able to make a good approximation of the image by solving a convex optimisation program.

III. SAR Generative Model with Phase Errors

Since SAR systems are a coherent imaging system, the round trip propagation delay to a reference position in the scene must be estimated at each position along the aperture. In spotlight mode SAR this reference point is the scene centre. Errors in this estimate, which can be due to a non-idealised propagation medium or inaccuracies in the inertial navigation system, introduce unknown phase errors into the acquired data. If not corrected, phase errors can degrade and produce distortions in the reconstructed image.

If we consider a simplified spotlight-mode SAR system after dechirp-on-receive, adding a delay error τ_{e} at each aperture position produces the following discretized system model [19],

$$\begin{aligned} y_{kl} = & \mathrm{e}^{j\phi_{kl}} \\ & \sum_{m=1}^{M} \sum_{n=1}^{N} x_{mn} \exp\left\{ -j \left(\frac{2u_{mnk}}{c} - \tau_0 \right) \left(\omega_0 + 2\alpha \left((l-1) T_s - \frac{T}{2} \right) \right) \right\} \end{aligned} \tag{6}$$

where, $\boldsymbol{Y} = \{y_{kl}\} \in \mathbb{C}^{M' \times N'}$ is the phase history, $\boldsymbol{X} = \{x_{mn}\} \in \mathbb{C}^{M \times N}$ are the scene reflectivities, $\{\phi_{kl}\} = (\omega_0 \tau_{\mathrm{e}_k} - \alpha \tau_{\mathrm{e}_k}^2) + 2\alpha\tau_{\mathrm{e}_k}((l-1)T_s - T/2) \in \mathbb{C}^{M' \times N'}$ are the phase errors which result from the delay errors, $\{u_{mnk}\} \in \mathbb{R}^{M \times N \times M'}$ are the distances between each element in the scene and each aperture position, c is the speed of light, τ_0 is the true propagation delay to the scene centre, T_s is the range sampling period, 2α is the chirp rate, ω_0 is the carrier frequency and T is the chirp period. If we neglect the effects of the linear phase term, which is done in most systems because this term usually has only a minor effect on the reconstructed image quality, the discrete SAR observation model with phase errors becomes:

$$\boldsymbol{Y} = \operatorname{diag}\left\{ \mathrm{e}^{j\boldsymbol{\phi}} \right\} h\left(\boldsymbol{X} \right), \tag{7}$$

where, $h : \mathbb{C}^{M \times N} \rightarrow \mathbb{C}^{M' \times N'}$ is a linear map that models the ideal SAR observation model (the summation in Eq. (6)) and

$$\phi_k = \omega_0 \tau_{\mathrm{e}_k} - \alpha \tau_{\mathrm{e}_k}^2 \tag{8}$$

are the phase errors.

Clearly, without further assumptions, the problem of recovering $\boldsymbol{\phi}$ and $\boldsymbol{X}$ from $\boldsymbol{Y}$ is ill-posed if $M' = M$ and $N' = N$, since there are only MN equations and $M(N+1)$ unknowns.

IV. CS with Post-processing Autofocus

Most post-processing autofocus methods make a far-field and small aperture angle approximation in the SAR acquisition model [19], i.e. the image was formed using a separable two-dimensional imaging method such as range-Doppler imaging [20]. Under the separable approximation and assuming we sample at exactly the Nyquist rate in range and cross range, the system can be modelled as the following LHS and RHS matrix multiplication:

$$\boldsymbol{Y} = \operatorname{diag}\left\{ \mathrm{e}^{j\boldsymbol{\phi}} \right\} \boldsymbol{AXB}, \tag{9}$$

where,

$$\begin{aligned} a_{mn} = \exp\{ & -j(2\pi(m-1)(n-1)/M - \\ & (m-1)\pi - (n-1)\pi + M\pi/2)\} \end{aligned}$$

and

$$\begin{aligned} b_{mn} = \exp\{ & -j(2\pi(m-1)(n-1)/N - \\ & (m-1)(2\pi\omega_0/2\alpha T - \pi) - \\ & (n-1)\pi + N\pi/2 - 2\omega_0 L/c)\} \end{aligned}$$

are the elements of the cross-range matrix $\boldsymbol{A} \in \mathbb{C}^{M\times M}$ and the range matrix $\boldsymbol{B} \in \mathbb{C}^{N\times N}$, respectively, where, L is the scene radius.

Since, $\boldsymbol{A}$ is essentially a Fourier matrix, we can rewrite the observation model in Eq. (9) as $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{\Psi}\boldsymbol{X}\boldsymbol{B}$, where, $\boldsymbol{\Psi}$ is a *circulant* matrix which may be viewed as a filter in cross-range direction for each range bin.

When fully-sampled, recovering $\boldsymbol{\Psi X}$ from $\boldsymbol{Y}$ is straight forward because $\boldsymbol{A}$ and $\boldsymbol{B}$ are invertible. Post-processing autofocus algorithms then recover $\boldsymbol{X}$ from the filtered image $\boldsymbol{\Psi X}$ using a signal model for $\boldsymbol{\Psi}$ and/or $\boldsymbol{X}$.

When $\boldsymbol{Y}$ is under-sampled in either range or cross-range the observation model will be:

$$\boldsymbol{Y'} = \boldsymbol{A}\boldsymbol{\Psi}\boldsymbol{X}\boldsymbol{B'} \tag{10}$$

or

$$\boldsymbol{Y'} = \boldsymbol{A'}\boldsymbol{\Psi}\boldsymbol{X}\boldsymbol{B}, \tag{11}$$

where, $\boldsymbol{A'} \in \mathbb{C}^{M'\times M}$ is a $M' < M$ row subset of $\boldsymbol{A}$ and $\boldsymbol{B'} \in \mathbb{C}^{N\times N'}$ is a $N' < N$ column subset of $\boldsymbol{B}$. With this model, unlike in the fully-sampled situation, $\boldsymbol{A'}$ and $\boldsymbol{B'}$ are not invertible.

An estimate of $\boldsymbol{\Psi X}$ can be reconstructed by solving Eq. (3). CS results can then be used to analyse the expected reconstruction quality of this estimate. If the under-sampling is random in cross-range, the reconstruction of the filtered image is stable, in the sense that the columns of the recovered filtered image $\widetilde{\boldsymbol{\Psi X}}$ satisfy Eq. (4), if the number of cross-range samples is greater than $\mathcal{O}(K\log^5(M))$. For a fixed K, the reconstruction error is dependent on the additive noise and the K-term approximations of the columns of the filtered image. Larger phases errors will make these K-term approximations worse and therefore increase the error in the reconstructed filtered image.

With an estimate of the filtered image the restructured image can be recovered by applying a standard post-processing autofocus technique. The resulting image is given by

$$\tilde{\boldsymbol{X}} = \tilde{\boldsymbol{\Psi}}^{-1}\widetilde{\boldsymbol{\Psi X}}, \tag{12}$$

where, $\tilde{\boldsymbol{\Psi}}^{-1} = \boldsymbol{A}^{-1}\operatorname{diag}\left\{\mathrm{e}^{-j\tilde{\boldsymbol{\phi}}}\right\}\boldsymbol{A}$ is the phase error correction applied by the autofocus algorithm and $\tilde{\boldsymbol{\phi}}$ are the estimated phase errors by the chosen algorithm. If the estimated phase errors are the true phase errors then the error in the reconstructed image is given by

$$\boldsymbol{X} - \tilde{\boldsymbol{X}} = \boldsymbol{\Psi}^{-1}\boldsymbol{E} \tag{13}$$

where, $\boldsymbol{E} = \boldsymbol{\Psi X} - \widetilde{\boldsymbol{\Psi X}}$ is the error in the estimated filtered image. Therefore, even with knowledge of the true phase errors, the effect of correcting phase errors as a post-processing step can result in a significant error in the reconstructed image. For this reason, in most cases, post-processing autofocus methods are unsuitable for under-sampled SAR.

V. Uniqueness

It is well known that there are inherit ambiguities in the autofocus problem which prevent the problem having a unique solution. The formulation in Eq. (9) is known to be ambiguous to constant and linear phase errors [19].

A sparsity based necessary condition for the uniqueness of the autofocus problem can be given which is dependent on the observation model h and the signal model of the scene X. It is given as follows:

$$h\left(\tilde{\boldsymbol{X}}\right) = \operatorname{diag}\{\boldsymbol{d}\}h\left(\boldsymbol{X}\right) \iff \exists\beta \in \{\beta \in \mathbb{C} : |\beta| = 1\} : \tilde{\boldsymbol{X}} = \beta\boldsymbol{X}, \tag{14}$$

and

$$\forall(\tilde{\boldsymbol{X}}, \boldsymbol{X}, \boldsymbol{d}) \in \left\{\tilde{\boldsymbol{X}} \in \mathcal{X}, \boldsymbol{X} \in \mathcal{X}, \boldsymbol{d} \in \mathcal{D}\right\},$$

where,

$$\mathcal{X} = \left\{\boldsymbol{X} \in \mathbb{C}^{M\times N} : \|\boldsymbol{X}\|_0 \leq K\right\},$$

i.e. we know the scene has at most K scatters, and

$$\mathcal{D} = \left\{\boldsymbol{d} \in \mathbb{C}^{M'} : |d_m| = 1\right\}$$

is the set of all possible phase errors,

If Eq. (14) is satisfied then the problem is unique up to a scalar β multiplication of the true $\boldsymbol{X}$, i.e. $\tilde{\boldsymbol{X}} = \beta\boldsymbol{X}$, and the solutions are given by the following program:

$$\begin{aligned} \underset{\boldsymbol{X},\boldsymbol{d}}{\text{minimise}} \quad & \|\boldsymbol{X}\|_0 \\ \text{subject to} \quad & \operatorname{diag}\{\boldsymbol{d}\}\boldsymbol{Y} = h\left(\boldsymbol{X}\right) \\ & d_m^* d_m = 1,\ m = 1,\ldots,M, \end{aligned} \tag{15}$$

where, $\|.\|_0$ measures the number of non-zeros matrix elements.

Eq. (14) states that the phase error free observation model h must have the property that the phase history of a sparse image cannot be equal to a phase error corrupted phase history of a different sparse image.

In Appendix A, we give additional conditions for the uniqueness of the separable model where we have sub-sampling only in the cross-range direction.

VI. Sparse Reconstruction and Autofocus

In this section our goal is to design algorithms which perform sparse reconstruction and autofocus and are able to be solved or approximately solved in a polynomial time. To this end, the non-convex function $\|\boldsymbol{X}\|_0$ in Eq. (15) is replaced with its closest convex function $\|\boldsymbol{X}\|_1$ and the equality constraint is replaced with an inequality constraint that accommodates noise. This results in:

$$\begin{aligned} \underset{\boldsymbol{X},\boldsymbol{d}}{\text{minimise}} \quad & \|\boldsymbol{X}\|_1 \\ \text{subject to} \quad & \|\operatorname{diag}\{\boldsymbol{d}\}\boldsymbol{Y} - h(\boldsymbol{X})\|_{\mathrm{F}} \leq \sigma \\ & d_m^* d_m = 1,\ m = 1,\ldots,M. \end{aligned} \tag{16}$$

Even though our objective function is now convex, Eq. (16) is still non-convex because the inequality constraint is not linear and therefore does not define a convex feasible set.

In order to use gradient based methods, which are usually used in large scale problems such as SAR reconstruction, the objective must be smooth. Therefore it is convenient to exchange to the inequality constraint and the objective in Eq. (16) to form the equivalent program:

$$\begin{aligned} \underset{\boldsymbol{X},\boldsymbol{d}}{\text{minimise}} \quad & \left\|\text{diag}\left\{\boldsymbol{d}\right\}\boldsymbol{Y}-h(\boldsymbol{X})\right\|_{\text{F}}^2 \\ \text{subject to} \quad & \left\|\boldsymbol{X}\right\|_1 \leq \tau \\ & d_m^* d_m = 1,\ m=1,\ldots,M. \end{aligned} \tag{17}$$

Note, there is a one-to-one map, $\gamma : \sigma \rightarrow \tau$ if $0 \leq \sigma \leq \left\|\boldsymbol{Y}\right\|_{\text{F}}$. Even though the problem is still non-convex, importantly, in each set of variables $\boldsymbol{X}$ and $\boldsymbol{d}$ –with the other fixed– we have a unique solution. This observation allows us to use a block relaxation type method which can be used to approximate the solution and has been found to be effective in the related problem of dictionary learning [21].

Block relaxation methods approximately solve Eq. (17) by iteratively solving the problem based on a single parameter block, $\boldsymbol{X}$ or $\boldsymbol{d}$, at a time.

A. Minimisation based on $\boldsymbol{X}$

Consider Eq (17) when $\boldsymbol{d}$ is fixed, i.e.

$$\begin{aligned} \underset{\boldsymbol{X}}{\text{minimise}} \quad & f(\boldsymbol{X},\boldsymbol{d}) \\ \text{subject to} \quad & \left\|\boldsymbol{X}\right\|_1 \leq \tau, \end{aligned} \tag{18}$$

where,

$$f(\boldsymbol{X},\boldsymbol{d}) = \left\|\text{diag}\left\{\boldsymbol{d}\right\}\boldsymbol{Y}-h(\boldsymbol{X})\right\|_{\text{F}}^2. \tag{19}$$

A method used for solving Eq. (18) is a technique known as "majorisation minimisation". This technique replaces the objective function with a majorising surrogate function which is much easier to solve. A function g is said to majorise f if $f(\omega) \leq g(\omega,\xi)$ and $f(\omega) = g(\omega,\omega), \forall \omega$ and $\xi \in \Upsilon$, where, Υ is the parameter space. A surrogate function can be derived for (19) by expanding it as a Taylor series and bounding its curvature ($d^2 f$) [21]. This surrogate function is:

$$\begin{aligned} g(\boldsymbol{X},\boldsymbol{X}^{\ddagger},\boldsymbol{d}) = & \left\|\text{diag}\left\{\boldsymbol{d}\right\}\boldsymbol{Y}-h(\boldsymbol{X})\right\|_{\text{F}}^2 - \\ & \left\|h(\boldsymbol{X})-h(\boldsymbol{X}^{\ddagger})\right\|_{\text{F}}^2 + \\ & L_X\left\|\boldsymbol{X}-\boldsymbol{X}^{\ddagger}\right\|_{\text{F}}^2, \end{aligned} \tag{20}$$

where, $L_X > \left\|h\right\|_2^2$. Replacing the objective function with its surrogate function, Eq. (18) becomes

$$\begin{aligned} \underset{\boldsymbol{X},\boldsymbol{X}^{\ddagger}}{\text{minimise}} \quad & g\left(\boldsymbol{X},\boldsymbol{X}^{\ddagger},\boldsymbol{d}\right) \\ \text{subject to} \quad & \left\|\boldsymbol{X}\right\|_1 \leq \tau, \end{aligned} \tag{21}$$

which is a minimisation based on $\boldsymbol{X}$ and a surrogate parameter vector $\boldsymbol{X}^{\ddagger}$. In this program, if $\boldsymbol{X}$ is fixed, the minimum of Eq. (21) occurs at $\boldsymbol{X}^{\ddagger} = \boldsymbol{X}$ and if $\boldsymbol{X}^{\ddagger}$ is fixed the minimum occurs at

$$\begin{aligned} \underset{\boldsymbol{X}}{\text{minimise}} \quad & \left\|\boldsymbol{X}-\boldsymbol{C}\right\|_{\text{F}} \\ \text{subject to} \quad & \left\|\boldsymbol{X}\right\|_1 \leq \tau, \end{aligned} \tag{22}$$

where, $\boldsymbol{C} = \boldsymbol{X}^{\ddagger} + \frac{1}{L_X} h^{\text{H}}(\text{diag}\left\{\boldsymbol{d}\right\}\boldsymbol{Y} - h(\boldsymbol{X}^{\ddagger}))$. The solution of Eq. (22) is the projection of $\boldsymbol{C}$ onto an ℓ_1 ball with a radius of τ. There are efficient methods to exactly compute this projection [22].

By minimising Eq. (21) based on either $\boldsymbol{X}^{\ddagger}$ and $\boldsymbol{X}$ in an alternating fashion, $\boldsymbol{X}^{\ddagger}$ and $\boldsymbol{X}$ will converge to the solution of Eq. (18) [23]. In practice, a feasible L_X can determined using a backtracking line-search.

B. Minimisation based on $\boldsymbol{d}$

Consider Eq. (17) when $\boldsymbol{X}$ is fixed, which (ignoring constant terms) is given by:

$$\begin{aligned} \underset{\boldsymbol{d}}{\text{minimise}} \quad & \text{tr}\left\{-2\,\text{Re}\left\{\text{diag}\left\{\boldsymbol{d}^{\text{H}}\right\}h(\boldsymbol{X})\boldsymbol{Y}^{\text{H}}\right\}\right\} \\ \text{subject to} \quad & d_m^* d_m = 1,\ m=1,\ldots,M. \end{aligned} \tag{23}$$

The unique solution of Eq. (23) can be found analytically by,

$$\boldsymbol{d} = e^{j\angle \text{diag}\left\{h(\boldsymbol{X})\boldsymbol{Y}^{\text{H}}\right\}}. \tag{24}$$

C. Non-convex Block Relaxation

A block relaxation of Eq. (17) is produced by solving Eq. (18) and Eq. (23) in an alternating fashion which is described in the following pseudo code:

Algorithm 1 $\mathcal{A}(\boldsymbol{X},\boldsymbol{d})$

Output: $\boldsymbol{X},\boldsymbol{d}$

repeat

$\boldsymbol{X}^{\ddagger} \leftarrow \boldsymbol{X}$

$\boldsymbol{X} \leftarrow \mathcal{D}(\boldsymbol{X},\boldsymbol{d})$

$\boldsymbol{d}^{\ddagger} \leftarrow \boldsymbol{d}$

$\boldsymbol{d} \leftarrow e^{j\angle \text{diag}\left\{h(\boldsymbol{X})\boldsymbol{Y}^{\text{H}}\right\}}$

until $\left\|\boldsymbol{X}-\boldsymbol{X}^{\ddagger}\right\|_{\text{F}}\left\|\boldsymbol{X}^{\ddagger}\right\|_{\text{F}}^{-1} <$ threshold$\wedge\left\|\boldsymbol{d}-\boldsymbol{d}^{\ddagger}\right\|_2\left\|\boldsymbol{d}^{\ddagger}\right\|_2^{-1} <$ threshold

Where, $\mathcal{D}$ solves Eq. (18). The approaches used in [12], [24] and [25] are of this form. These types of methods are stable, assuming we can solve $\mathcal{D}$, i.e. we exactly solve Eq. (18) at each iteration. In practical algorithms where only an approximate solution at each iteration is obtained, no stability analysis exists.

Another way to create a block relaxation is to use the surrogate parameter $\boldsymbol{X}^{\ddagger}$ as an additional parameter block, i.e.

$$\begin{aligned} \underset{\boldsymbol{X},\boldsymbol{X}^{\ddagger},\boldsymbol{d}}{\text{minimise}} \quad & g\left(\boldsymbol{X},\boldsymbol{X}^{\ddagger},\boldsymbol{d}\right) \\ \text{subject to} \quad & \left\|\boldsymbol{X}\right\|_1 \leq \tau \\ & d_m^* d_m = 1,\ m=1,\ldots,M. \end{aligned} \tag{25}$$

For this relaxation, as long as Eq. (25) is always solved based on $\boldsymbol{X}^{\ddagger}$ after solving based on $\boldsymbol{X}$ the solution for each sub-problem is easily commutable and the complete

algorithm is known to be stable and guaranteed to converge to an accumulation point or a connected set of accumulation points, see [21, Proposition B.3]. The pseudo code for this algorithm, when phase minimisation occurs at each iteration, is as follows:

Algorithm 2 $\mathcal{B}(\boldsymbol{X}, \boldsymbol{d})$

Initialise: $L_x > \|h\|_{\mathrm{F}}^2$
Output: $\boldsymbol{X}, \boldsymbol{d}$
repeat
$\boldsymbol{X}^{\ddagger} \leftarrow \boldsymbol{X}$
$\boldsymbol{C} \leftarrow \boldsymbol{X}^{\ddagger} + \frac{1}{L} h^{\mathrm{H}} \left(\mathrm{diag}\left\{\boldsymbol{d}\right\} \boldsymbol{Y} - h\left(\boldsymbol{X}^{\ddagger}\right)\right)$
$\boldsymbol{X} \leftarrow \mathcal{P}_{\tau}(\boldsymbol{C})$
$\boldsymbol{d}^{\ddagger} \leftarrow \boldsymbol{d}$
$\boldsymbol{d} \leftarrow e^{j \angle \mathrm{diag}\left\{h(\boldsymbol{X}) \boldsymbol{Y}^{\mathrm{H}}\right\}}$
until $\|\boldsymbol{X} - \boldsymbol{X}^{\ddagger}\|_{\mathrm{F}} \|\boldsymbol{X}^{\ddagger}\|_{\mathrm{F}}^{-1} <$ threshold $\wedge \|\boldsymbol{d} - \boldsymbol{d}^{\ddagger}\|_2 \|\boldsymbol{d}^{\ddagger}\|_2^{-1} <$ threshold

Where, $\mathcal{P}_{\tau}(\boldsymbol{C})$ projects $\boldsymbol{C}$ onto an ℓ_1 ball with a radius of τ. It is interesting to note that this algorithm can be seen as a generalisation of Algorithm 1. An additional benefit of Algorithm 2 is that it is likely to converge faster than Algorithm 1. This is because Algorithm 1 will likely oscillate around the optimum path.

VII. Experimental Results

In these experiments we investigate the performance of Algorithm 1 and Algorithm 2 using under-sampled phase histories that contain phase errors.

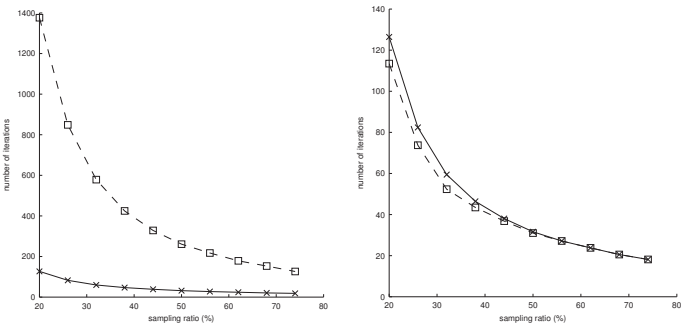
A. Quantitative Performance

In the first experiment we investigate the empirical convergence rate and reconstruction performance of Algorithm 1 and Algorithm 2. In order to easily compare with post-processing autofocus techniques, we consider the separable model, Eq. (9). In this experiment the scene consists of a small number of constant amplitude point targets randomly placed in the scene. The under-sampling consists of selecting a random subset of the fully-sampled synthetic aperture. Two different phase errors were consider: quadratic phase errors $\phi_m = \gamma((m-1)/M)^2$ which model platform velocity measurement errors and normally distributed phase errors $\phi_m = \mathcal{N}(0, \gamma^2)$. The parameters for the synthetic model are in Table. I.

TABLE I
SAR system parameters for synthetic experiments

parameter	value
carrier frequency (ω_0)	$2\pi \times 10 \times 10^9$ rad/s
chirp bandwidth ($2\alpha T$)	$2\pi \times 150 \times 10^6$ rad/s
scene radius (L)	50 m
number of targets	20
signal to noise ratio	0 dB

1) Convergence: In this experiment we compare the number of iterations it takes Algorithm 1 and Algorithm 2 to reach the stopping criterion when the threshold is 10^{-6}. In order to fairly compare the two algorithms we compute the operation $\mathcal{D}$ in Algorithm 1 using the "majorisation minimisation" method from Section VI-A. We also define the number of iterations in each algorithm to be the total number of times the gradient of the objective function has to been computed with respect to $\boldsymbol{X}$. We select this definition because the main computational cost of both algorithms is consumed by computing this gradient, therefore, the iterations count will closely relate to the algorithm's execution time. We choose to show the results for normally distributed phase errors with $\gamma = 10$. This is because the type and magnitude of phase errors was found to have only a minor effect on the results.



(a) (b)

Fig. 2. Comparison of empirical convergence rates: (a) '□' Algorithm 1 and '×' Algorithm 2. (b) '□' Algorithm 2 with continuation and '×' Algorithm 2.

As expected Fig. 2(a) shows that Algorithm 2 requires many less iterations than Algorithm 1. This will likely be due to the minimisation path of Algorithm 1 oscillating around the optimal minimisation path.

A technique known as continuation has been found to be useful for increasing the numerical convergence rate of ℓ_1 sparse recovery algorithms when there is no phase errors [26]. Continuation involves varying the value of τ during the iterations of the algorithm. The motivation for this technique is based on the observation that the convergence rate depends on τ. The smaller than value of τ, the faster the algorithm will converge. Therefore, a method of continuation is to start with a small value of τ and increases its value in the following iterations until it reaches the desired final value.

In order to further improved the convergence rate of our algorithm we experimented with a continuation scheme. Although we did not see any singularity in the modified algorithm with this setting, the convergence and stability would need to be proved in the future.

In this simulation we used a continuation scheme that involved changing τ during the first I iterations by the rule $\tau_i = i\tau/I$ for $i = 1, \ldots, I$. The selection of a "good" I depends on the under-sampling so we used the following values of I for each under-sampling percentage.

TABLE II
Continuation parameters

sampling ratio (%)	20	26	32	38	44	50	56	62	68	74
I	30	20	10	5	3	2	1	1	1	1

Fig. 2(b) shows a small improvement in performance when continuation is used. Another method for reducing the required number of iterations would be to use a more aggressive step size, similar to what is used is other iterative ℓ_1 sparse recovery algorithms. Using this type of step size, the stability of the algorithm cannot be guaranteed but in practise it may also be useful.

2) Reconstruction Error: In order to assess the image reconstruction performance of the autofocus methods we define an image quality metric. Since the autofocus problem is ambiguous to scalar multiplication by $\beta \in \{\beta \in \mathbb{C} : |\beta| = 1\}$ and cyclic permutation, we define a metric that is immune to these ambiguities. We will refer to this metric as relative SNR and define it as:

$$\underset{\beta,n}{\text{minimise}} \left\{ 10 \log_{10} \left(\frac{\|\boldsymbol{X}\|_{\text{F}}^2}{\left\| \tilde{\boldsymbol{X}} - \beta \boldsymbol{P}^n \boldsymbol{X} \right\|_{\text{F}}^2} \right) \right\},$$

where, $n \in \mathbb{Z}$ and

$$\boldsymbol{P} = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix}.$$

Fig. 4 shows the reconstruction performance of sparse recovery with post-processing autofocus and Algorithm 2 with different phase errors. We do not show the results of Algorithm 1 because the results are virtually identical to that of Algorithm 2. The magnitude of the corresponding filters for each of the phase errors, the rows of $\boldsymbol{\Psi}$, are shown in Fig. 3.

To provide an empirical upper-bound, we also show the reconstruction performance that can be achieved with oracle knowledge of the phase errors and also the locations of the targets, we refer to this as the oracle reconstruction. The oracle reconstruction recovers an image as follows: it first corrects the phase errors in the phase history such that it has no phase errors. It then uses the known location of the targets to perform a least squares (LS) estimate of the target reflectivities. This problem is overdetermined since there are K reflectivities and $M'N > K$ measurements.

The sparse recovery with post-processing autofocus is performed as is described in Section IV. Firstly, an ℓ_1-norm spectral projected gradient (SPG) method [23] is used to recover the filtered image $\widetilde{\boldsymbol{\Psi}\boldsymbol{X}}$ then the image is recovered from the filtered image using the reference phase errors $\tilde{\boldsymbol{\phi}}$.

These reference phase errors are selected slightly differently for the two different types of phase errors. When the phase errors are quadratic, $\tilde{\boldsymbol{\phi}}$ is selected to be equal to the true phase errors, even for the phases associated with unobserved measurements. This is because the CS reconstruction tends to approximate the blurred image. However, when the phase errors are random, $\tilde{\boldsymbol{\phi}}$ is selected to be equal to the true phase error at the M' indices corresponding to the observed aperture measurements and 0 at all indices corresponding to unobserved aperture measurements. The reason for this difference is

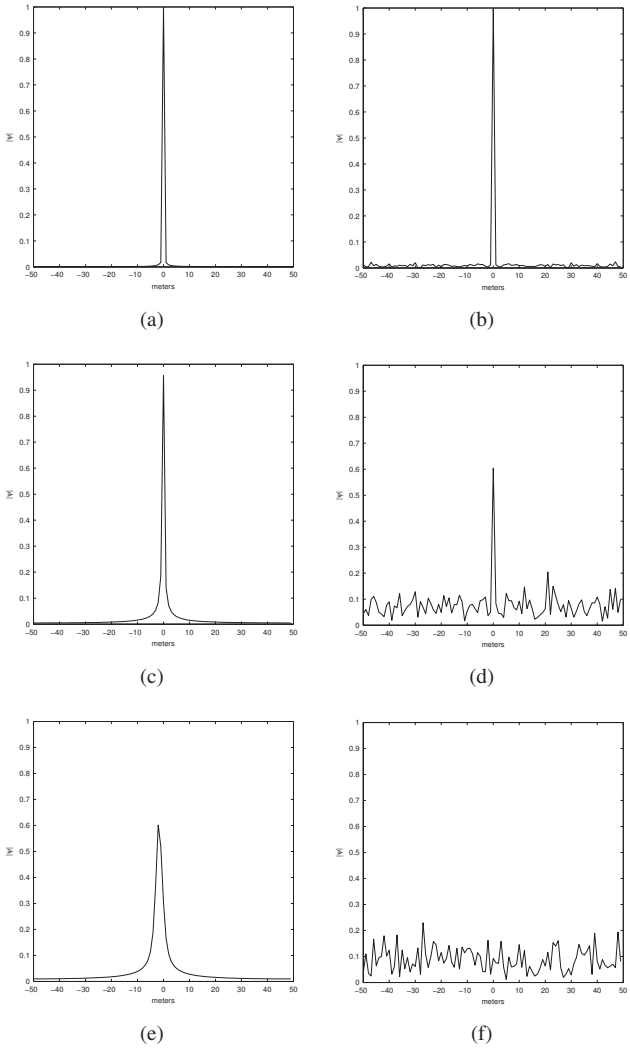


Fig. 3. Phase error induced filters: the rows of Ψ for quadratic and random phase errors with different γ. Quadratic: (a) $\gamma = 0.1$ (c) $\gamma = 1$ (e) $\gamma = 10$. Random: (b) $\gamma = 0.1$ (d) $\gamma = 1$ (f) $\gamma = 10$.

because for the random phase errors, unlike quadratic phase errors, each of the phase errors are independent. This means that the phase errors associated with the unobserved measurements have no effect on the sparse recovery algorithm, therefore, it doesn't make sense to use them in $\tilde{\boldsymbol{\Psi}}^{-1}$. In comparison, setting the unobserved phase errors to zero slightly increases the reconstruction performance.

To understand the results of Fig. 4, the sources of errors in the reconstructed image should be considered. The three sources of errors in a reconstructed image are the additive noise, the under-sampling and the phase errors. For the oracle reconstruction the only source of error is due to the additive noise. The performance degrades with the sampling ratio because the denoising effect of the LS estimate degrades as the ratio $M'N/K$ decreases. For the sparse recovery with autofocus the sparse recovery will try to minimise the errors associated with the under-sampling and also will implicitly denoise. This process will be more successful if the filtered image is approximately sparse. The post-processing autofocus

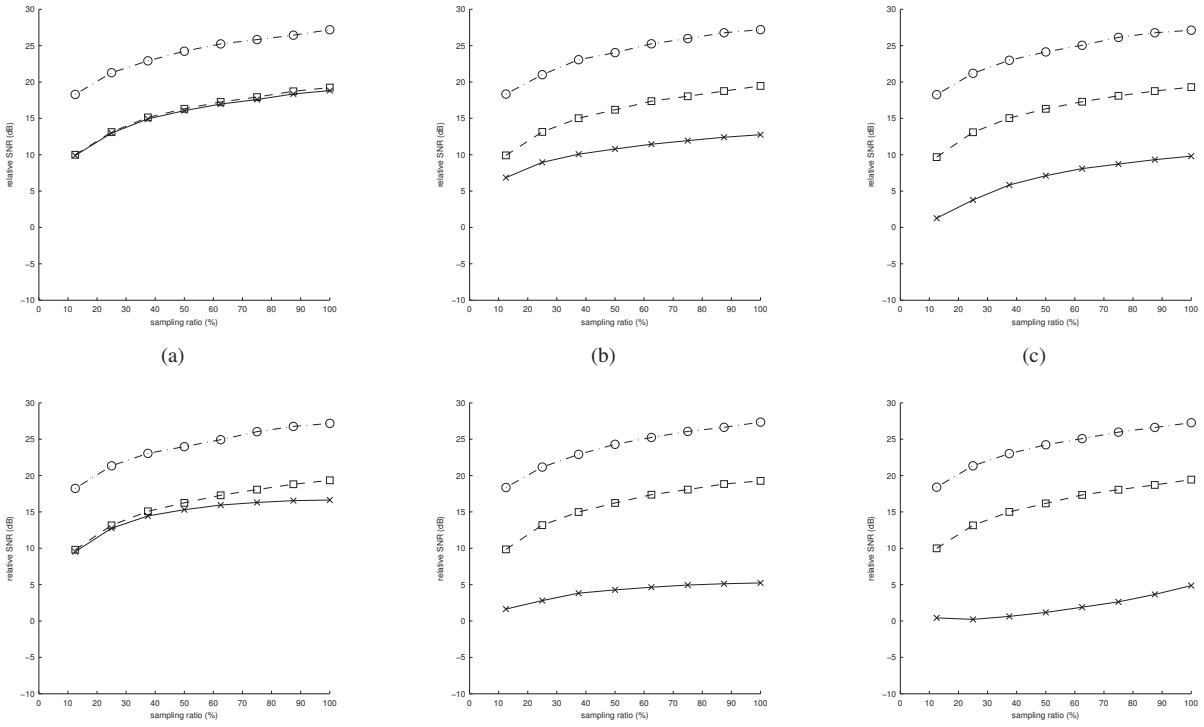


Fig. 4. Reconstruction performance versus under-sampling ratio: '$\circ$' oracle reconstruction, '$\square$' Algorithm 2 and '$\times$' sparse recovery and post-processing autofocus with reference phase errors. Quadratic: (a) $\gamma = 0.1$ (b) $\gamma = 1$ (c) $\gamma = 10$. Random: (d) $\gamma = 0.1$ (e) $\gamma = 1$ (f) $\gamma = 10$.

will then try to reduce the errors associated with the phase errors. Algorithm 2 also minimises the effect of under-sampling and phase errors and implicitly denoises.

As predicted in Section IV, as the phase errors increase, the performance of sparse recovery with post-processing autofocus decreases. It is also interesting to note that this methods performance is better for quadratic phase errors than for random phase errors. This is because the filter corresponding to the quadratic phase errors is approximately sparse while than the filter corresponding to the random phase errors is not. Hence, the sparse recovery for quadratic phase errors is more effective at reducing the errors associated with the undersampling and the additive noise.

The performance of Algorithm 2, which is in contrast to the performance of sparse recovery with post-processing autofocus, is consistently good for both types of phase errors. In fact, it achieves a performance, even with large phase errors, that is similar to a sparse recovery without phase errors. The SNR gap between the performance of Algorithm 2 and the oracle reconstruction is primarily due to the shrinkage effects of ℓ_1-minimisation. This gap could potential be reduced by an additional procedure known as "debiasing" [27].

B. Qualitative Performance

In these experiments we wish to show that the presented algorithm works on realistic simulations of our two motivating scenarios, i.e. multifunction and UWB SAR. The scene used in both simulated scenarios consists of four point targets which reflect back an equal amount of energy. Fig. 5 is a block diagram which illustrates the basic elements used to create the simulated phase histories. Firstly, the analog signal that would be received at each aperture position is simulated by summing scaled and delayed versions of the transmitted chirp where the scaling and delay correspond to the reflectivity and the signal travel time for each point target. For each position an additional delay is added to the analog received signal to model system inaccuracies. Each analog signal is then dechirped and IF filtered which simulates the analog receiver in a dechirp-on-receive system. Finally the analog to digital sampling is simulated by down sampling the signals to a sample rate proportional to the IF bandwidth and the Residual Video Phase (RVP) term is removed.

1) UWB SAR: As mentioned previously, under sampling occurs in a UWB SAR system when notches are introduced into the transmitted chirp in order to avoid interference with other users. In this simulation we used a notched linear frequency chirp which had a spectral density that is given in Fig. 6. The chirp contains five notches which equate to a nulling of approximately 20% of the chirp spectrum.

The other parameters of the simulation are given below.

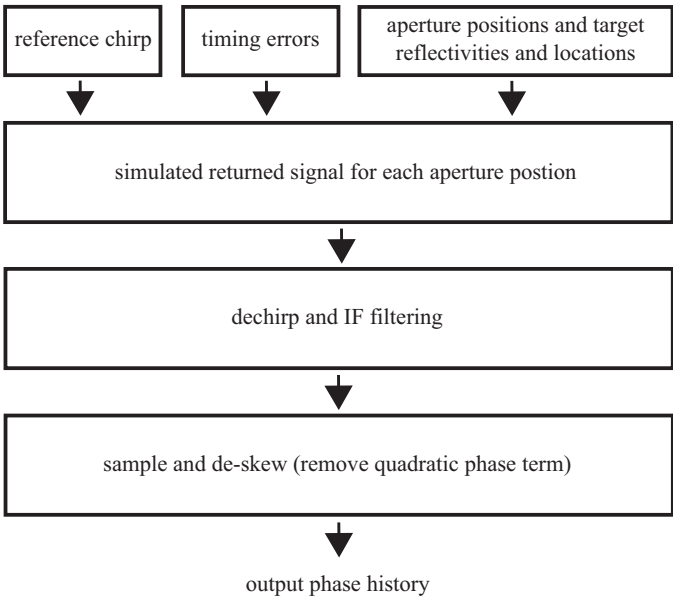


Fig. 5. Block diagram for generating a simulated phase history.

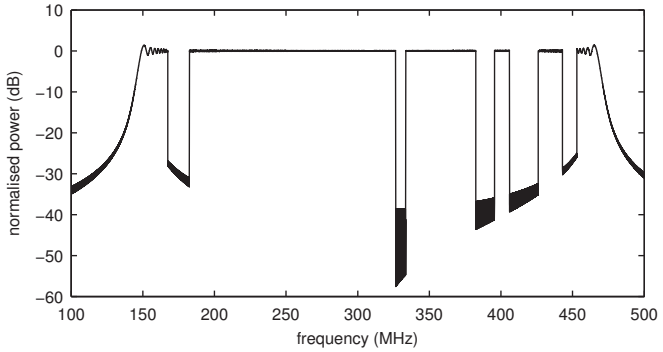


Fig. 6. Power spectral density of notched linear frequency chirp.

TABLE III
SYSTEM PARAMETERS FOR SIMULATED UWB SAR

parameter	value
carrier frequency (ω_0)	$2\pi \times 308 \times 10^6$ rad/s
chirp bandwidth ($2\alpha T$)	$2\pi \times 324 \times 10^6$ rad/s
IF bandwidth	$2\pi \times 20 \times 10^6$ rad/s
altitude	7000 m
stand-off distance	7000 m
aperture length	7000 m
number of aperture samples	200
scene radius (L)	75 m
number of targets	4
signal to noise ratio	0 dB
timing errors	$\mathcal{N}(0, 80 \times 10^-11)$ s

2) Multifunction SAR: In this simulation a randomly under-sampled aperture of an X-band SAR system is used to simulate a multifunction SAR system. The phase history contains a 50% random subset of the fully-sampled aperture. The other parameters of the simulation are given below.

TABLE IV
SYSTEM PARAMETERS FOR SIMULATED MULTIFUNCTION SAR

parameter	value
carrier frequency (ω_0)	$2\pi \times 10 \times 10^9$ rad/s
chirp bandwidth ($2\alpha T$)	$2\pi \times 600 \times 10^6$ rad/s
IF bandwidth	$2\pi \times 30 \times 10^6$ rad/s
altitude	7000 m
stand-off distance	7000 m
aperture length	250 m
number of aperture samples	300
scene radius (L)	75 m
number of targets	4
signal to noise ratio	0 dB
timing errors	$\mathcal{N}(0, 2.5 \times 10^-11)$ s

For both scenarios, three SAR images where formed using different reconstruction methods. One image in each scenarios was generated using filtered back-projection without any form of autofocus. Another was generated using 20 iterations of an ℓ_1-norm SPG method again without any form of autofocus. The last image was created using 20 iterations of the modified Algorithm 2 which uses continuation with $I = 15$. The final value of τ was selected to be the sum of the absolute values of the target reflectivities. However, the reconstruction performance was found to be not particularly dependent on this parameter. In a real system a suitable τ could be selected with only a coarse degree of parameter tuning. In the iterative reconstruction algorithms both the observation model and its adjoint ($h(.)$ and $h^{\mathrm{H}}(.)$) are computed using the fast (re/back)-projection algorithms from [25].

The resulting images from both simulation scenarios are contained in Fig. 7 and Fig. 8. It should be noted that these images have had been padded with zeroes in the spatial Fourier domain to make the images twice the size of the reconstructed images. This is done to more clearly display the point targets which are sometimes unable to be clearly viewed when they consist of only a single or a small number of non-zero pixels.

Fig. 7(a) and Fig.8(a) demonstrate the adverse effects of phase errors and under-sampling. The side-lobes of the four targets contain a large amount of energy which deteriorates the SAR image quality. The images in Fig. 7(c) and Fig.8(c) which were produced using an ℓ_1 sparse recovery algorithm have an improved visual quality over the previous images due to the sparsity promoting algorithm. However, due to the model inaccuracies there are a large number of non-zeros pixels that may be mistaken for additional targets. Finally, Fig. 7(e) and Fig.8(e) show the results of Algorithm 2. In these images the energy from each target is highly concentrated around the target locations. It is clear that in these scenarios, Algorithm 2 can produce a visually improved SAR image, with a rectangular spatial Fourier support and a sparse number of point targets, from a phase history that is under-sampled and contains model inaccuracies.

VIII. CONCLUSION

We have investigated the effects of phase errors on an under-sampled SAR system. We have shown that post-processing autofocus algorithms are typically unsuitable when there is under-sampling and a sparse reconstruction method is em-

ployed. Instead, phase errors should be corrected during the image reconstruction.

We have proposed a new algorithm that corrects phase errors within the image reconstruction algorithm. Algorithm 2, which is an algorithmically stable generalisation of a recently proposed non-convex sparsity based autofocus method, performs consistently well for a variety of phase errors and under-sampling ratios and was found empirically to converge in a much smaller number of iterations.

We have also demonstrated through additional realistic simulations that Algorithm 2 could be used in practical non-standard SAR image reconstruction systems to produce sparse SAR images from under-sampled phase histories which contain model inaccuracies.

Although we have concluded that post-processing autofocus algorithms are typically unsuitable for under-sampled SAR, there may be some instances where they may warrant further consideration. In the scenario where the under-sampling is only in the range dimension, for example the UWB scenario, an ℓ_1-based sparse recovery algorithm could be used to perform range compression and then a standard reconstruction method could be used to form the final image which could then be autofocused using a standard post-processing algorithm. CS theory suggests that this will be sub-optimal, however, this type of method may be justified as a means of reducing complexity. Further research into autofocus and image reconstruction algorithms for under-sampled data where there are specific system constraints could be an avenue for future research.

APPENDIX A

Using ideas from the dictionary learning literature [28] we can define a set of sufficient conditions for the uniqueness of $\boldsymbol{\phi}$ and $\boldsymbol{X}$ given $\boldsymbol{Y'} = \mathrm{diag}\{\boldsymbol{d'}\}\boldsymbol{A'XB}$. These conditions are as follows:

1) the spark condition: any $2K_{\boldsymbol{X}}$ columns of $\boldsymbol{A'}$ are linearly independent
2) the columns of $\boldsymbol{X}$ have exactly $K_{\boldsymbol{X}}$ non-zero elements
3) for each of the $\binom{M}{K_{\boldsymbol{X}}}$ possible $K_{\boldsymbol{X}}$-sparse supports, there are at least $K_{\boldsymbol{X}}+1$ columns of $\boldsymbol{X}$
4) any $K_{\boldsymbol{X}}+1$ columns of $\boldsymbol{X}$ which share the same support, span a k-dimensional space
5) any $K_{\boldsymbol{X}}+1$ columns of X, which have different supports, span a $(K_{\boldsymbol{X}}+1)$-dimensional space

Proposition 1 (see [28, Theorem 3]): If the above conditions hold then there is a unique $\tilde{\boldsymbol{X}}$ which satisfies $\boldsymbol{Y'} = \mathrm{diag}\left\{\tilde{\boldsymbol{d'}}\right\}\boldsymbol{A'}\tilde{\boldsymbol{X}}\boldsymbol{B}$. Where uniqueness is up to a unit magnitude scalar β and a circular permutation $\boldsymbol{P}^n$ of the true $\boldsymbol{X}$, i.e. $\tilde{\boldsymbol{X}} = \beta\boldsymbol{P}^n\boldsymbol{X}$

As is the case in dictionary learning, the richness condition 3 is completely unrealistic for compressively sampled SAR. However, this condition is only sufficient and is likely to be very pessimistic. It should also be noted that recovering the unique solution involves solving Eq. (15) which requires combinatorial many operations to solve and is unsuitable for practical problems that involve noise.

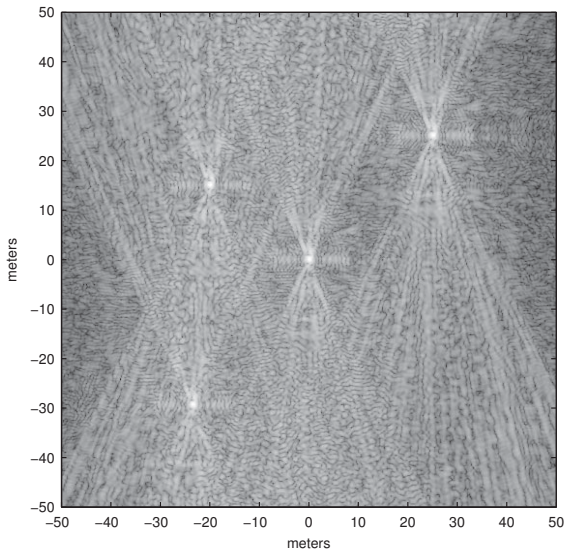
ACKNOWLEDGEMENT

This work forms part of the UK University Defence Research Collaboration in Signal Processing and was supported by EPSRC/DSTL grants [EP/H012370/1, EP/J015180/1 and EP/K014277/1].

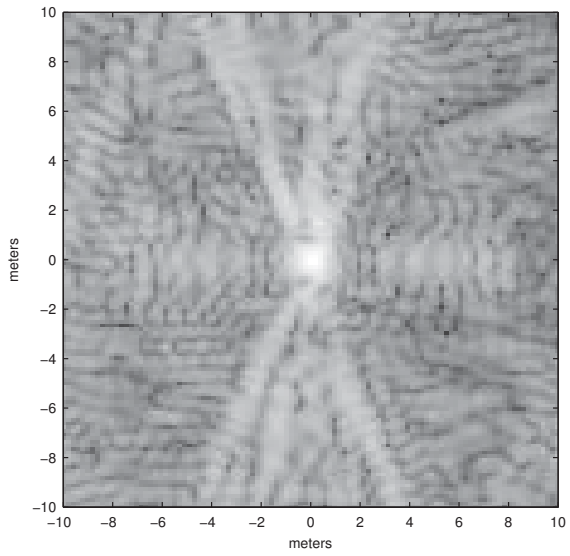
REFERENCES

[1] J. Salzman, D. Akamine, R. Lefevre, and J. Kirk, J.C., "Interrupted synthetic aperture radar (SAR)," in *Proc. of IEEE Radar Conf.*, May 2001, pp. 117–122.
[2] D. Vu, L. Xu, M. Xue, and J. Li, "Nonparametric missing sample spectral analysis and its applications to interrupted SAR," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 1, pp. 1–14, Feb. 2012.
[3] M. Davis, *Foliage penetration radar*, 1st ed. SciTech Publishing, 2011.
[4] X. Luo, L. Ulander, J. Askne, G. Smith, and P.-O. Frolind, "RFI suppression in ultra-wideband SAR systems using LMS filters in frequency domain," *Elect. Lett.*, vol. 37, no. 4, pp. 241–243, Feb. 2001.
[5] V. Patel, G. R. Easley, D. Healy, and R. Chellappa, "Compressed synthetic aperture radar," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 244–254, 2010.
[6] L. Potter, E. Ertin, J. Parker, and M. Cetin, "Sparsity and compressed sensing in radar imaging," *Proc. IEEE*, vol. 98, no. 6, pp. 1006–1020, 2010.
[7] L. Nguyen and T. Do, "Recovery of missing spectral information in ultra-wideband synthetic aperture radar (SAR) data," in *Proc. of IEEE Radar Conf.*, May 2012, pp. 253–256.
[8] I. Stojanovic, W. Karl, and L. Novak, "Reconstruction of interrupted SAR imagery for persistent area surveillance change detection," in *Proc. of SPIE*, vol. 8394, Apr. 2012.
[9] W. Carrara, R. Goodman, and R. Majewski, *Spotlight-mode synthetic aperture radar: signal processing algorithms*, 1st ed. Artech House, 1995.
[10] D. Wahl, P. Eichel, D. Ghiglia, and C. Jakowatz Jr, "Phase gradient autofocus-a robust tool for high resolution SAR phase correction," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 30, no. 3, pp. 827–835, Jul. 1994.
[11] R. Morrison Jr, M. Do, and D. Munson Jr, "MCA: A multichannel approach to SAR autofocus," *IEEE Trans. Image Process.*, vol. 18, no. 4, pp. 840–853, Apr. 2009.
[12] N. Onhon and M. Cetin, "A sparsity-driven approach for joint SAR imaging and phase error correction," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2075–2088, Apr. 2012.
[13] E. Candes, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *Commun. Pure, Applied Math.*, vol. 66, no. 8, p. 12411274, Aug. 2013.
[14] X. Li and V. Voroninski, "Sparse signal recovery from quadratic measurements via convex programming," 2012, preprint, arXiv:1209.4785 [cs.IT].
[15] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, 2006.
[16] S. Foucart, "Sparse recovery algorithms: sufficient conditions in terms of restricted isometry constants," in *Springer Proc. in Math.*, vol. 13, 2010, pp. 65–77.
[17] E. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.
[18] M. Rudelson and R. Vershynin, "On sparse reconstruction from Fourier and Gaussian measurements," *Comm. Pure Appl. Math*, vol. 61, no. 8, pp. 1025–1045, Aug. 2008.
[19] C. Jakowatz Jr, D. Wahl, P. Eichel, D. Ghiglia, and P. Thompson, *Spotlight-mode synthetic aperture radar: a signal processing approach*, 4th ed. Kluwer Academic Publishers, 1999.
[20] M. Soumekh, *Synthetic aperture radar signal processing with MATLAB algorithms*, 1st ed. John Wiley & Sons, 1999.
[21] M. Yaghoobi, T. Blumensath, and M. Davies, "Dictionary learning for sparse approximations with the majorization method," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2178–2191, Jun. 2009.
[22] J. Duchi, S. Shalev-Shwartz, Y. Singer, and T. Chandra, "Efficient projections onto the ℓ_1-ball for learning in high dimensions," in *Proc. of ICML*, 2008, pp. 272–279.
[23] E. Birgin, J. Martínez, and M. Raydan, "Nonmonotone spectral projected gradient methods on convex sets," *SIAM J. Optim.*, vol. 10, no. 4, pp. 1196–1211, 2000.

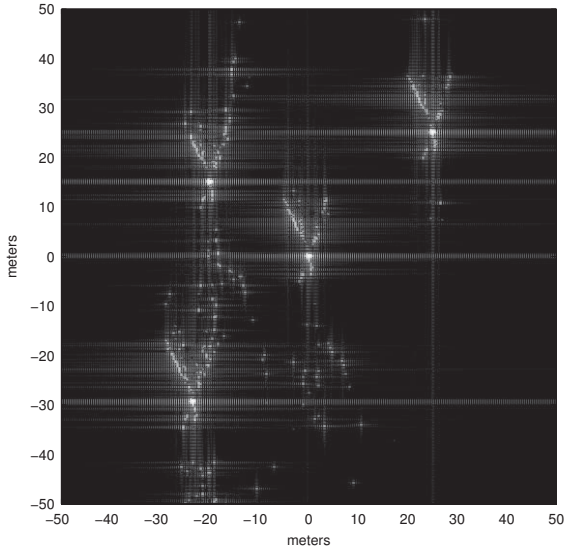
[24] M. Onhon and M. Cetin, "A nonquadratic regularization-based technique for joint SAR imaging and model error correction," in *Proc. of SPIE*, vol. 7337, Apr. 2009.
[25] S. Kelly, G. Rilling, M. Davies, and B. Mulgrew, "Iterative image formation using fast (re/back)-projection for spotlight-mode SAR," in *Proc. of IEEE Radar Conf.*, May 2011, pp. 835–840.
[26] E. Hale, W. Yin, and Y. Zhang, "Tr07-07: A fixed-point continuation method for l1-regularized minimization with applications to compressed sensing," Rice University, Tech. Rep., 2007.
[27] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 586–597, 2007.
[28] M. Aharon, M. Elad, and A. Bruckstein, "On the uniqueness of overcomplete dictionaries, and a practical way to retrieve them," *J. Linear Algebra Appl.*, vol. 416, no. 1, pp. 48–67, Jul. 2006.



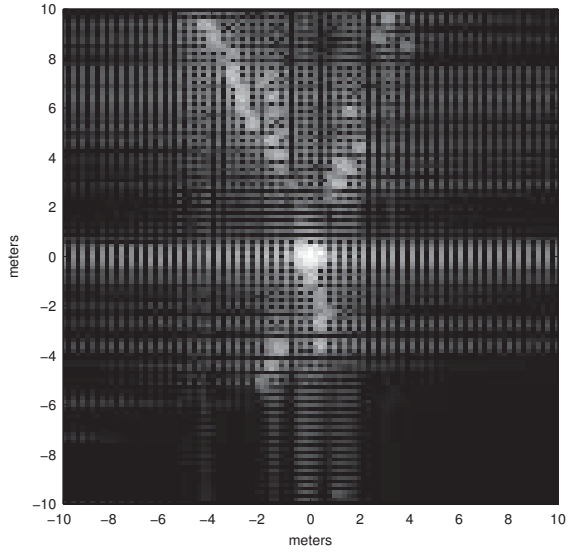
(a)



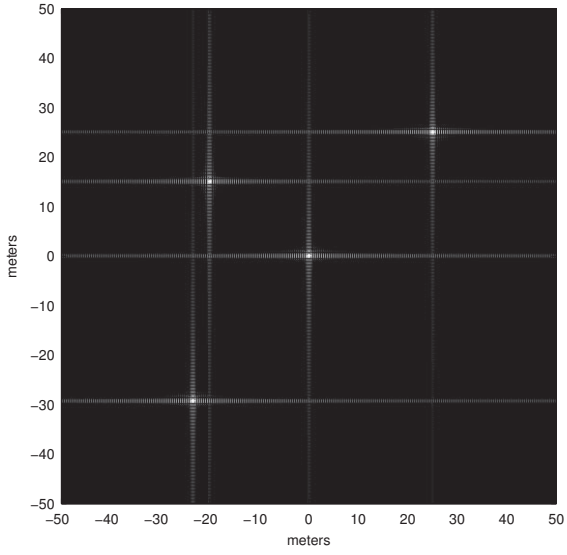
(b)



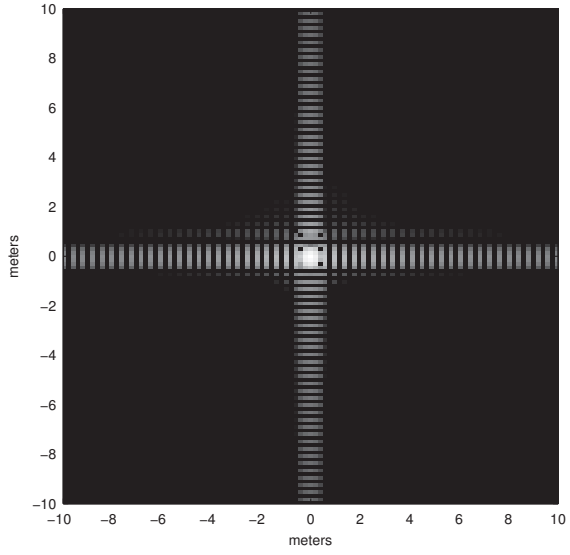
(c)



(d)

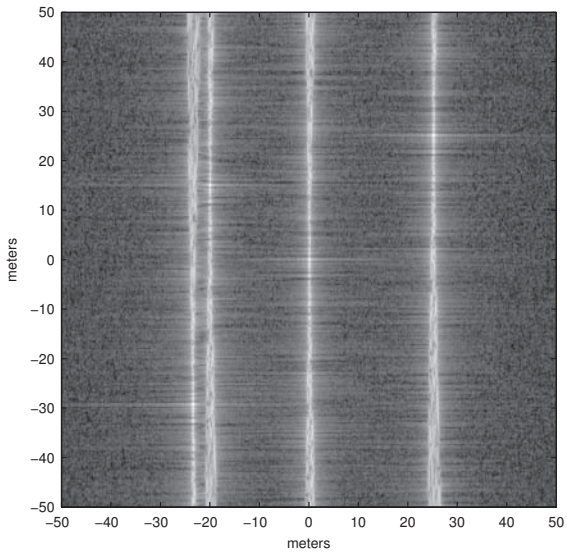


(e)

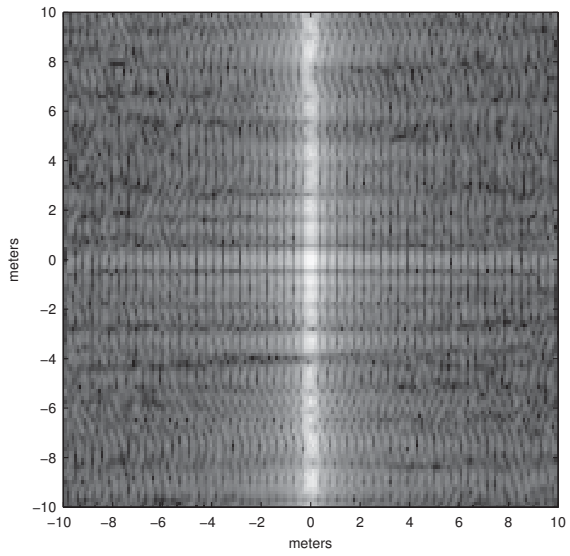


(f)

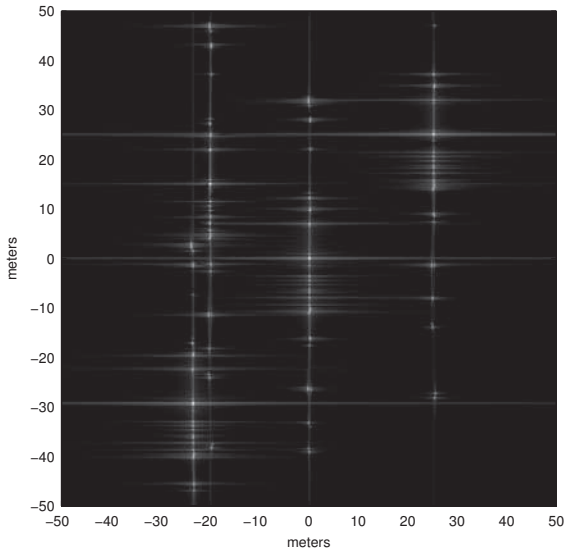
Fig. 7. UWB SAR image reconstructions: (a) was reconstructed using filtered back-projection, (c) was reconstructed using an ℓ_1-norm SPG method and (e) was reconstructed using Algorithm 2. (b), (d) and (f) are a zoomed in view of (a), (c) and (e) around the origin, respectively.



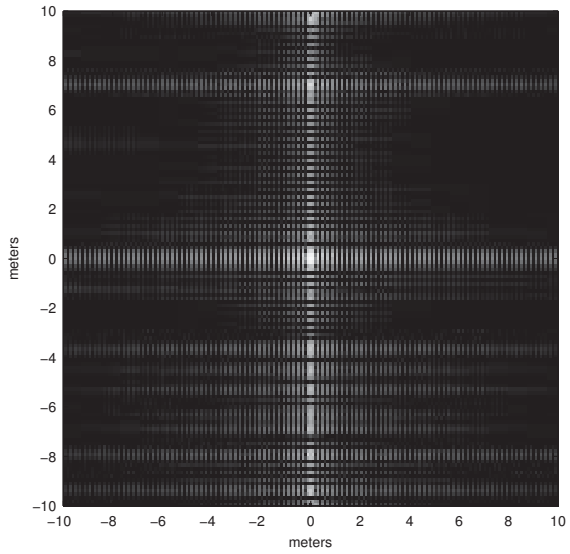
(a)



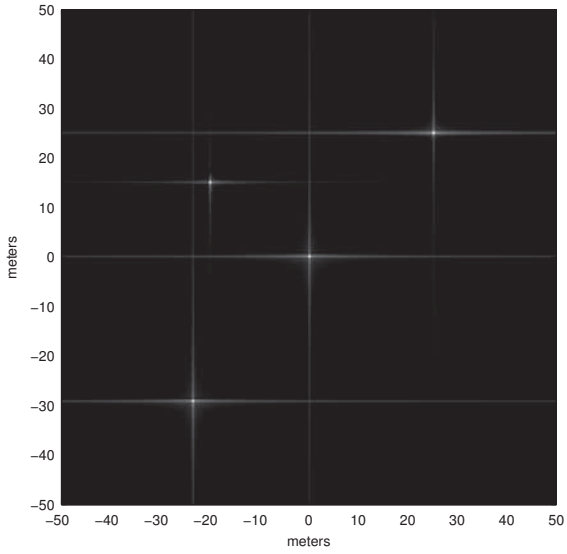
(b)



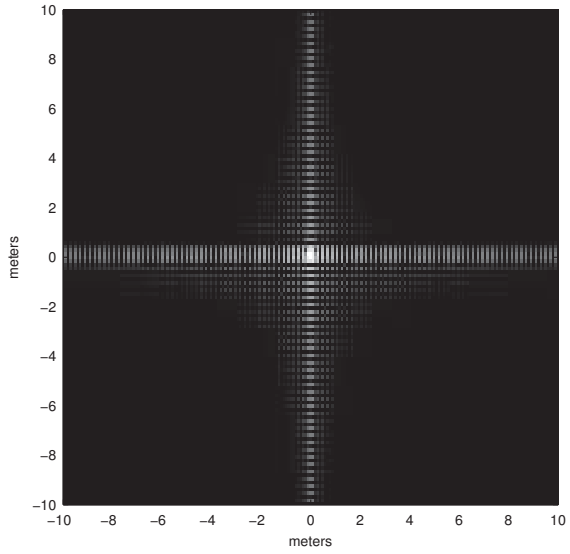
(c)



(d)



(e)



(f)

Fig. 8. Multifunction SAR image reconstructions: (a) was reconstructed using filtered back-projection, (c) was reconstructed using an ℓ_1-norm SPG method and (e) was reconstructed using Algorithm 2. (b), (d) and (f) are a zoomed in view of (a), (c) and (e) around the origin, respectively.